

Uma ontologia pode ser definida como uma especificação formal e explícita de uma conceitualização compartilhada por um domínio de interesse, onde formal significa que a ontologia deve ser interpretável por computador e aceita por um grupo ou comunidade da área que a ontologia modela. Esta pesquisa apresenta um trabalho em andamento na área de de aprendizado de ontologias a partir de texto, focando na identificação de conceitos e relações. Foi utilizada uma proposta baseada em (a) métricas de grafos para identificar conceitos, e (b) análise do texto com os conceitos encontrados para obter relações. A abordagem proposta resulta em uma ontologia com 535 conceitos e 29419 relações, e é dividida em quatro partes: aquisição de conceitos, aquisição de sinônimos, aquisição de relações taxonômicas e relações não-taxonômicas. A aquisição de termos consiste em encontrar automaticamente palavras que representem conceitos de um domínio. A identificação de sinônimos visa a aquisição semântica de variantes de termos, ou seja, encontrar entre os termos de um texto aqueles que compartilham funções semânticas. A extração de relações taxonomicas busca identificar uma organização hierárquica entre os conceitos, enquanto a extração de relações não-axonômicas tem por objetivo descobrir relações entre conceitos (por exemplo, relações entre sintomas, doenças e drogas) através da análise de textos. O aprendizado semi-automático de ontologias tem apresentado resultados promissores, e capazes de diminuir o custo da construção de ontologias. Esse trabalho contribui para avançar as pesquisas que necessitam de informação estruturada, tais como, agentes conversacionais e sistemas de perguntas e respostas. Especialmente em língua portuguesa que possui recursos escassos na área de processamento de linguagem.